On Cantor's first proof of uncountability

Peng Kuan 彭宽 *titang78@gmail.com*

17 February 2016

Abstract: Discussion about Cantor's first proof using the next-interval-function, potential and actual infinity.

1. Cantor's first proof

Cantor's first proof of the uncountability of real numbers is the first rigorous demonstration of the notion of uncountability. Countable sets can be put into a list indexed with natural numbers. If a set cannot be listed, then, it has more members than the set of natural numbers and is uncountable. Cantor's first proof is a proof by contradiction. First, he supposes that all real numbers are listed in any order by the list $X=(x_1, x_2, x_3 \dots)$. Then, a real number out of this list is found by using a series of intervals, contradicting that X lists all real numbers.

The first 2 elements of the original list X are x_1 and x_2, which define the first interval $[a_1, b_1]$ of the series:

$$\begin{aligned} &\text{If } x_1<x_2, \text{ then } [a_1, b_1] = [x_1, x_2] \\ &\text{If } x_2<x_1, \text{ then } [a_1, b_1] = [x_2, x_1] \end{aligned} \qquad (1)$$

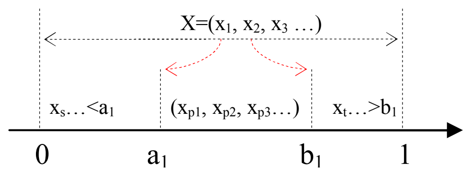

Figure 1

The interval $[a_1, b_1]$ divides the remaining $(x_3, x_4, x_5 \dots)$ into 3 parts. The elements $x_{p1}, x_{p2}, x_{p3} \dots$ are within $[a_1, b_1]$ and form the next list X_1. The elements $x_s\dots$ and $x_t\dots$ are outside $[a_1, b_1]$ and are excluded. Figure 1 shows this process. The next list $X_1=(x_{p1}, x_{p2}, x_{p3} \dots)$ keeps the indexes of the original list. For example, the element x_{1234} of the original list X is the element x_{p1} and the original indexes is $p_1=1234$.

For creating the series of intervals, the above process is recursively applied on succeeding intervals. The i^{th} list $X_i=(x_{i1}, x_{i2}, x_{i3} \dots)$ is inside the i^{th} interval $[a_i, b_i]$. The $(i+1)^{th}$ interval $[a_{i+1}, b_{i+1}]$ is defined by the first 2 elements x_{i1} and x_{i2}. Then, the $(i+1)^{th}$ list $X_{i+1}=(x_{r1}, x_{r2}, x_{r3} \dots)$ is constructed with those elements of $(x_{i3}, x_{i4}, x_{i5} \dots)$ which are within $[a_{i+1}, b_{i+1}]$. This process is performed by the next-interval-function *nI*:

$$\begin{gathered} \{[a_{i+1}, b_{i+1}], X_{i+1}\} = nI\,(X_i) \\ \text{Such that :} \\ [a_{i+1}, b_{i+1}] = [x_{i1}, x_{i2}] \text{ if } x_{i1} < x_{i2} \\ [a_{i+1}, b_{i+1}] = [x_{i2}, x_{i1}] \text{ if } x_{i2} < x_{i1} \\ X_{i+1} = (x_{r1}, x_{r2}, x_{r3} \dots) \in \,]\, a_{i+1}, b_{i+1} \,[\end{gathered} \qquad (2)$$

The elements in all lists are so arranged that their original index increases, that is, $i_1< i_2< i_3<\dots$ The original indexes of a_{i+1} and b_{i+1} are i_1 and i_2 which are smaller than i_3 and those in the $(i+1)^{th}$ list X_{i+1} which is selected from $(x_{i3}, x_{i4}, x_{i5} \dots)$:

$$\text{Original indexes of } a_{i+1} \text{ and } b_{i+1} < i_3 \leq \text{Original indexes in } X_{i+1} \qquad (3)$$

The bounds of the next interval a_{i+2} and b_{i+2} are the first 2 elements of X_{i+1} and thus, have original indexes bigger than that of a_{i+1} and b_{i+1}. So, the original indexes of interval bounds increase with i:

$$\begin{gathered} \text{Original indexes of } a_{i+1} \text{ and } b_{i+1} < \text{Original indexes of } a_{i+2} \text{ and } b_{i+2} \\ \Rightarrow \text{Original indexes of } a_i \text{ and } b_i \to \infty \text{ when } i \to \infty \end{gathered} \qquad (4)$$

For proving that the real numbers in the interval [0, 1] are not countable, we suppose that the original list $X=(x_1, x_2, x_3 \ldots)$ lists all real numbers in the interval [0, 1]. Then we create the series of intervals by applying the next-interval-function *nI* on X to generate the first interval and list $\{[a_1,b_1],X_1\}$, then, on X_1 to generate the second interval and list $\{ [a_2, b_2], X_2\}$ and so on:

$$\begin{aligned}
&\{ [a_1, b_1], X_1\} = nI\,(X)\\
&\{ [a_2, b_2], X_2\} = nI\,(X_1\})\\
&\ldots\\
&\{ [a_n, b_n], X_n\} = nI\,(X_{n-1}\})
\end{aligned} \tag{5}$$

This recursive process generates the series of n nested intervals, $[a_1, b_1]$, $[a_2, b_2]$, …, $[a_n, b_n]$, where each interval includes the next one:

$$[a_1, b_1] \supset [a_2, b_2] \supset \ldots \supset [a_n, b_n] \tag{6}$$

This way, the bounds of intervals, $a_1, a_2, \ldots, a_n$ and $b_1, b_2, \ldots, b_n$ form an increasing sequence:

$$\begin{aligned}
&[a_i, b_i] \supset [a_{i+1}, b_{i+1}] \text{ for all } i<\infty\\
&\Rightarrow a_i< a_{i+1}< b_{i+1}<b_i \text{ for all } i<\infty\\
&\Rightarrow a_1< a_2< a_3<\ldots< a_n < b_n<\ldots< b_3< b_2< b_1
\end{aligned} \tag{7}$$

When $n \to \infty$, a_n increases while staying smaller than b_n; on the other hand, b_n decreases while staying bigger than a_n. So, a_n and b_n tends to a limit z:

$$z = \lim_{n\to\infty} a_n = \lim_{n\to\infty} b_n \Rightarrow a_n < z < b_n \text{ for all } n < \infty \tag{8}$$

According to equation **(3)**, $[a_i, b_i]$ contains only elements having bigger original indexes than that of a_i and b_i. So, for whatever original index m, the element x_m of X is excluded by those intervals $[a_n, b_n]$ whose original indexes are bigger than m. According to equation (4), when $n\to\infty$ the original indexes of a_n and b_n increase indefinitely and all elements of X will be excluded by $[a_n, b_n]$. If a number is never excluded when $n\to\infty$, then, it is not an element of X.

According to equation (8) z is inside $[a_n, b_n]$ for all n. Thus z will never be excluded and is not an element of the original list X, contradicting the supposition that X lists all real numbers of the interval [0, 1]. In consequence, the real numbers in the interval [0, 1] are not countable.

This is Cantor's first proof. Its reasoning is limited to potential infinity because n never reaches infinity.

2. Under the logic of potential infinity

Potential infinity does not allow $n=\infty$. Under this logic, the number of terms of an infinite sequence is $n<\infty$ while $n\to\infty$. So, the sequences $(a_1, a_2, a_3, \ldots, a_n)$ and $(b_1, b_2, b_3, \ldots, b_n)$ are finite. The last terms a_n and b_n are not identical but $a_n < b_n$ for all $n<\infty$.

Because $b_n - a_n > 0$, the interval $[a_n, b_n]$ is a segment of the real line with non-zero length and contains always infinity of real numbers. Because a_n and b_n are the bounds of $[a_n, b_n]$ within which limits of a_n and b_n have no sense, the real number z does not exist. In fact, the limits of a_n and b_n exist only on the scale of [0, 1] because $b_n-a_n \to 0$ only with respect to 1. To understand this fact we just need to change the variable x to y such that:

$$y = \frac{x - a_n}{b_n - a_n} \tag{9}$$

The interval $[a_n, b_n]$ for the variable x is the interval [0, 1] for the variable y. Creating a list of real numbers in $[a_n, b_n]$ will create the list $Y=(y_1, y_2, y_3\ldots)$ in [0, 1]. Let $Y=(y_1, y_2, y_3\ldots)$ to equal $X=(x_1, x_2, x_3\ldots)$ and we are back to the starting point: the same list Y = X in the same interval [0, 1]. For all n, $[a_n, b_n]$ contains the entire original list X.

No outlying real number z exists, no contradiction is found. So, under the logic of potential infinity Cantor's first proof is not valid.

Because Cantor's first proof is within the logic of potential infinity, my demonstration could normally end here. But mathematics is the art of always going beyond. So, let us see what does Cantor's first proof looks like under the logic of actual infinity.

3. Under the logic of actual infinity

Now n is allowed to reach infinity. To get a sense of the influence of $n=\infty$, let us see the sequence of decimal numbers that converges to π:

$\pi_1=3.1$	$\pi_2=3.14$	$\pi_3=3.141$	…	π_∞

Table 1

The index $n=1, 2, 3\ldots\infty$ is also the number of decimals of each term. Naturally, π_∞'s number of decimals is infinity and $\pi_\infty=\pi$, meaning that the ∞^{th} term of the sequence equals the limit. One may disagree and contend that the ∞^{th} term should not exist. If so, we are back under the logic of potential infinity with $n<\infty$. If one argues that no term of the sequence should equal the limit, then, $\pi_\infty\neq\pi$ and the sequence π_n does not converge to π but to π_∞. The logic of actual infinity imposes that the ∞^{th} term exists and equals the limit of the sequence.

So, the final interval $[a_\infty, b_\infty]$ exists and $a_\infty=b_\infty=z$. The bounds of the n^{th} interval $[a_n, b_n]$ are elements of the original list X. For $n=\infty$, $[a_\infty, b_\infty] = [a_n, b_n]$, implying that a_∞, b_∞ and z are elements of X too.

When the series of intervals is allowed to become infinite, z is well defined but is an element of the original list X. So, Cantor's first proof cannot find the necessary contradiction even under the logic of actual infinity and is invalid.

4. About uncountability

<<On the uncountability of the power set of ℕ>> shows that the proof of the uncountability of the power set of ℕ has no contradiction. <<Hidden assumption of the diagonal argument>> shows that the diagonal argument proof of the uncountability of real numbers has no contradiction either. Now, the contradiction of Cantor's first proof vanishes too. The 3 pillars of uncountability seem to fail all. The inevitable question is: Is uncountability real?